

#8

ENTERED

OIPE

RAW SEQUENCE LISTING

DATE: 07/26/2001

PATENT APPLICATION: US/09/788,268

TIME: 14:45:04

Input Set: A:\010262.txt

Output Set: N:\CRF3\07262001\I788268.raw

```

3 <110> APPLICANT: Jarvik, Jonathan W.
5 <120> TITLE OF INVENTION: Methods and Products for Peptide-Based DNA
6   Sequence Identification and Analysis
8 <130> FILE REFERENCE: 2087 010262
10 <140> CURRENT APPLICATION NUMBER: US 09/788,268
11 <141> CURRENT FILING DATE: 2001-02-16
13 <150> PRIOR APPLICATION NUMBER: PCT/US99/30104
14 <151> PRIOR FILING DATE: 1999-12-16
16 <150> PRIOR APPLICATION NUMBER: US 60/182,816
17 <151> PRIOR FILING DATE: 2000-02-16
19 <150> PRIOR APPLICATION NUMBER: US 60/189,310
20 <151> PRIOR FILING DATE: 2000-03-14
22 <160> NUMBER OF SEQ ID NOS: 28
24 <170> SOFTWARE: Microsoft Word 97 SR-2
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 6
28 <212> TYPE: PRT
29 <213> ORGANISM: Artificial Sequence
31 <220> FEATURE:
32 <223> OTHER INFORMATION: Example of sequence made up entirely of six-codon amino acids
34 <400> SEQUENCE: 1
35 Leu Arg Arg Leu Leu Arg
36   1           5
38 <210> SEQ ID NO: 2
39 <211> LENGTH: 6
40 <212> TYPE: PRT
41 <213> ORGANISM: Artificial Sequence
43 <220> FEATURE:
44 <223> OTHER INFORMATION: Example of sequence made up entirely of one-codon amino acids
46 <400> SEQUENCE: 2
47 Met Trp Trp Met Met Trp
48   1           5
50 <210> SEQ ID NO: 3
51 <211> LENGTH: 100
52 <212> TYPE: DNA
53 <213> ORGANISM: Homo sapiens
55 <400> SEQUENCE: 3
56 gaattctttac acctcatact ttcccaagcc ccaactttct catctgaaaa tggtaatagt 60
58 atcatcctta catgtttaag gtcatgaatt gctatgtgta 100
60 <210> SEQ ID NO: 4
61 <211> LENGTH: 16
62 <212> TYPE: PRT
63 <213> ORGANISM: Homo sapiens
65 <400> SEQUENCE: 4
66 Thr Met Ile Thr Pro Ser Leu His Ala Cys Arg Ser Thr Leu Glu Asp
67   1           5           10           15
69 <210> SEQ ID NO: 5

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/788,268

DATE: 07/26/2001

TIME: 14:45:04

Input Set : A:\010262.txt

Output Set: N:\CRF3\07262001\I788268.raw

```

70 <211> LENGTH: 100
71 <212> TYPE: DNA
72 <213> ORGANISM: Homo sapiens
74 <400> SEQUENCE: 5
75 gaattcacat aaatcgcaaa tttttttttt cttcccagag ccattcaaaa ctctgtttgt 60
77 caaaggcctg tctgaggata ccactgaaga gacattaaag 100
79 <210> SEQ ID NO: 6
80 <211> LENGTH: 99
81 <212> TYPE: DNA
82 <213> ORGANISM: Homo sapiens
84 <400> SEQUENCE: 6
85 gaattctctt gggttttgtg gtgtgctaga cttaattacc catgaatgat tttgtcctct 60
87 tgagaaaatt tcaatagcac atctattagt gttttttat 99
89 <210> SEQ ID NO: 7
90 <211> LENGTH: 27
91 <212> TYPE: DNA
92 <213> ORGANISM: Artificial Sequence
94 <220> FEATURE:
95 <221> NAME/KEY: SITE
96 <222> LOCATION: (4)..(9)
97 <223> OTHER INFORMATION: Oligonucleotide primer containing EcoRI site
99 <400> SEQUENCE: 7
100 cccgaattca gcaggtaaaa atcaagg 27
102 <210> SEQ ID NO: 8
103 <211> LENGTH: 29
104 <212> TYPE: DNA
105 <213> ORGANISM: Artificial Sequence
107 <220> FEATURE:
108 <221> NAME/KEY: SITE
109 <222> LOCATION: (4)..(9)
110 <223> OTHER INFORMATION: Oligonucleotide primer containing EcoRI site
112 <400> SEQUENCE: 8
113 ggggaattct tactcttctc cactgctat 29
115 <210> SEQ ID NO: 9
116 <211> LENGTH: 24
117 <212> TYPE: DNA
118 <213> ORGANISM: Artificial Sequence
120 <220> FEATURE:
121 <223> OTHER INFORMATION: Nucleotide input sequence used to deonstrate computer
program
122 capabilities
124 <400> SEQUENCE: 9
125 caactagaag aggtaagaaa ctat 24
127 <210> SEQ ID NO: 10
128 <211> LENGTH: 8
129 <212> TYPE: PRT
130 <213> ORGANISM: Artificial Sequence
132 <220> FEATURE:
133 <223> OTHER INFORMATION: Computer program output of encoded peptides
135 <400> SEQUENCE: 10

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/788,268

DATE: 07/26/2001

TIME: 14:45:04

Input Set : A:\010262.txt

Output Set: N:\CRF3\07262001\I788268.raw

```

136 Gln Leu Glu Glu Val Arg Asn Tyr
138 <210> SEQ ID NO: 11
139 <211> LENGTH: 326
140 <212> TYPE: DNA
141 <213> ORGANISM: Homo sapiens
143 <220> FEATURE:
144 <221> NAME/KEY: exon
145 <222> LOCATION: (37).. (283)
147 <400> SEQUENCE: 11
148 gggaagccca tctccagctg tctgtttccc tttaagtcga atcaagagca acgtggatgg 60
149 gcggtacctg gtggacggcg tccctttcag ctgctgcaat cctagctcgc cacggccctg 120
150 catccagtat cagatcacca acaactcagc aactacagt tacgaccacc agacggagga 180
151 gctcaacctg tgggtgcgtg gctgcagggc tgccctgctg agctactaca gcagcctcat 240
152 gaactccatg ggtgtcgtca cgctcctcat ttggtctctc gaggtaggcc ctgggcagct 300
153 gggggtagag ggtaaggaga gcctcc 326
155 <210> SEQ ID NO: 12
156 <211> LENGTH: 36
157 <212> TYPE: DNA
158 <213> ORGANISM: Artificial sequence
160 <220> FEATURE:
161 <223> OTHER INFORMATION: Primer synthesized and used to PCR amplify rds/peripherin
exon 2
162 from an individual known to carry a wild type allele of
163 rds/peripherin.
165 <400> SEQUENCE: 12
166 ggcccgaat tctccagctg tctgtttccc tttaag 36
168 <210> SEQ ID NO: 13
169 <211> LENGTH: 37
170 <212> TYPE: DNA
171 <213> ORGANISM: Artificial sequence
173 <220> FEATURE:
174 <223> OTHER INFORMATION: Primer synthesized and used to PCR amplify rds/peripherin
exon 2
175 from an individual known to carry a wild type allele of
176 rds/peripherin.
178 <400> SEQUENCE: 13
179 aatttactcg agctaccccc agctgcccag ggcctac 37
181 <210> SEQ ID NO: 14
182 <211> LENGTH: 364
183 <212> TYPE: PRT
184 <213> ORGANISM: Artificial sequence
186 <220> FEATURE:
187 <223> OTHER INFORMATION: Fusion protein
189 <400> SEQUENCE: 14
190 Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro
191 1 5 10 15
192 Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu
193 20 25 30
194 Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu
195 35 40 45
196 Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/788,268

DATE: 07/26/2001

TIME: 14:45:04

Input Set : A:\010262.txt

Output Set: N:\CRF3\07262001\I788268.raw

```

197      50      55      60
198 Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn
199 65      70      75      80
200 Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu
201      85      90      95
202 Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser
203      100      105      110
204 Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu
205      115      120      125
206 Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn
207      130      135      140
208 Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp
209 145      150      155      160
210 Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu
211      165      170      175
212 Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr
213      180      185      190
214 Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala
215      195      200      205
216 Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Ile Glu Gly
217      210      215      220
218 Arg Gly Ile Gln Asp Leu Val Pro His Thr Thr Pro His His Thr Thr
219 225      230      235      240
220 Pro His His Thr Thr Pro His His Thr Thr Pro Gln Asp Leu Asn Ser
221      245      250      255
222 Pro Ala Val Cys Phe Pro Leu Ser Arg Ile Lys Ser Asn Val Asp Gly
223      260      265      270
224 Arg Tyr Leu Val Asp Gly Val Pro Phe Ser Cys Cys Asn Pro Ser Ser
225      275      280      285
226 Pro Arg Pro Cys Ile Gln Tyr Gln Ile Thr Asn Asn Ser Ala His Tyr
227      290      295      300
228 Ser Tyr Asp His Gln Thr Glu Glu Leu Asn Leu Trp Val Arg Gly Cys
229 305      310      315      320
230 Arg Ala Ala Leu Leu Ser Tyr Tyr Ser Ser Leu Met Asn Ser Met Gly
231      325      330      335
232 Val Val Thr Leu Leu Ile Trp Leu Phe Glu Val Gly Pro Gly Gln Leu
233      340      345      350
234 Gly Val Ala Arg Ser Ser Gly Arg Ile Val Thr Asp
235      355      360

```

237 <210> SEQ ID NO: 15

238 <211> LENGTH: 87

239 <212> TYPE: DNA

240 <213> ORGANISM: Artificial sequence

242 <220> FEATURE:

243 <221> NAME/KEY: misc_feature

244 <222> LOCATION: (35)..(37)

245 <223> OTHER INFORMATION: Upstream primer used to reamplify amplicons *ok*

246 Start codon at 35-37

248 <400> SEQUENCE: 15

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/788,268

DATE: 07/26/2001
 TIME: 14:45:04

Input Set : A:\010262.txt
 Output Set: N:\CRF3\07262001\I788268.raw

```

249 ggatcctaatac gactcact atagggagac caccatgcat caccatcatc accatcacca 60
250 ctctccagct gtctgtttcc cttaag 87
252 <210> SEQ ID NO: 16
253 <211> LENGTH: 35
254 <212> TYPE: DNA
255 <213> ORGANISM: Artificial sequence
257 <220> FEATURE:
258 <223> OTHER INFORMATION: Downstream primer used to reamplify amplicons
260 <400> SEQUENCE: 16
261 cttagtcatt atacccccag ctgccaggg cctac 35
263 <210> SEQ ID NO: 17
264 <211> LENGTH: 21
265 <212> TYPE: DNA
266 <213> ORGANISM: Homo sapiens
268 <400> SEQUENCE: 17
269 ttctctctct ctttatttta g 21
271 <210> SEQ ID NO: 18
272 <211> LENGTH: 24
273 <212> TYPE: DNA
274 <213> ORGANISM: Homo sapiens
276 <400> SEQUENCE: 18
277 actaaacaat gtacatgaac atac 24
279 <210> SEQ ID NO: 19
280 <211> LENGTH: 24
281 <212> TYPE: DNA
282 <213> ORGANISM: Homo sapiens
284 <220> FEATURE:
285 <221> NAME/KEY: variation
286 <222> LOCATION: (1)..(3)
288 <400> SEQUENCE: 19
289 tatttctcc tctctttatt ttag 24
291 <210> SEQ ID NO: 20
292 <211> LENGTH: 24
293 <212> TYPE: DNA
294 <213> ORGANISM: Homo sapiens
296 <400> SEQUENCE: 20
297 actaaacaat gtacatgaac atac 24
299 <210> SEQ ID NO: 21
300 <211> LENGTH: 27
301 <212> TYPE: DNA
302 <213> ORGANISM: Homo sapiens
304 <220> FEATURE:
305 <221> NAME/KEY: variation
306 <222> LOCATION: (1)..(6)
308 <400> SEQUENCE: 21
309 tattacttcc tcctctcttt atttttag ✓ 27
311 <210> SEQ ID NO: 22
312 <211> LENGTH: 24
313 <212> TYPE: DNA

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/788,268

DATE: 07/26/2001

TIME: 14:45:05

Input Set : A:\010262.txt

Output Set: N:\CRF3\07262001\I788268.raw